



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/057,523	01/22/2002	George M. White	2222.0820005	5053
26111 7590 02/16/2010 STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C. 1100 NEW YORK AVENUE, N.W. WASHINGTON, DC 20005				
EXAMINER				
LERNER, MARTIN				
ART UNIT		PAPER NUMBER		
2626				
MAIL DATE		DELIVERY MODE		
02/16/2010		PAPER		

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary

Application No.

10/057,523

Applicant(s)

WHITE ET AL.

Examiner

MARTIN LERNER

Art Unit

2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 12 January 2010.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 59, 61 to 66, 68 to 73, 75 to 79, 81, and 83 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 59, 61 to 66, 68 to 73, 75 to 79, 81, and 83 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date _____
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: _____

DETAILED ACTION

Claim Rejections - 35 USC § 112

1. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

2. Claims 59, 61 to 66, 68 to 73, 75 to 79, 81, and 83 are rejected under 35

U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Independent claims 59, 66, 73, 81, and 83 set forth the limitation of "the additional input", which is indefinite because it lacks antecedent basis. Moreover, although the claims recite receiving input, it is unclear when any "additional input" is received in relation to the original input.

3. The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

4. Claims 59, 61 to 66, 68 to 73, 75 to 79, 81, and 83 are rejected under 35

U.S.C. 112, first paragraph, as failing to comply with the written description requirement.

The claims contain subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventors, at the time the application was filed, had possession of the claimed invention.

Independent claims 59, 66, 73, 81, and 83 set forth the limitation of "updating the previously stored acoustic model based on the additional input", which limitation involves new matter. Applicants' Specification, corresponding to U.S. Patent Publication No. 2002/0072918, discloses at, ¶[0072], that the keywords recognizable by the distributed VUI system may be updated or modified according to the user's word choices, but does not disclose that the acoustic model may be updated or modified based on the user's word choices. The only significant disclosures of acoustic models from U.S. Patent Publication No. 2002/0072918 appear at ¶[0088], ¶[0090], ¶[0106] - ¶[0109], and ¶[0111] - ¶[0112]. However, the disclosure does not say that the acoustic models may be updated, but only, at best, that the acoustic models recognize speech based upon previous stored enunciations in reference voice templates at ¶[0088]. Although the Specification discloses that the keywords recognizable may be updated or modified by the user's word choices, there is nothing in the originally-filed Specification that says that the acoustic model may be updated. Thus, the limitation of "updating the previously stored acoustic model based on the additional input" introduces new matter.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

6. Claims 59, 63 to 64, 66, 70 to 71, 73, 77 to 78, 81, and 83 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Jacobs et al.* in view of *Miyazawa et al.*

Concerning independent claims 59, 81, and 83, *Jacobs et al.* discloses a distributed voice recognition system and method, comprising:

“a transceiver configured to receive input from the device via a communications network, wherein the input is the result of preliminary signal processing comprising keyword detection by the device prior to receipt of the input at the transceiver” – central communications center or base station 42 has receiver 46 and transmitter 50 (“a transceiver”), which receives features (“the input”) from portable phone 40 (“the device”) through a wireless network (“via a communications network”); a speech signal received at microphone 20 of portable phone 40 is provided to feature extraction element 22, which extracts relevant characteristics of the input speech (“preliminary signal processing”) (column 5, lines 21 to 56: Figure 2); in one embodiment, handset 100 (“the device”) recognizes a small number of simple, special voiced commands by local VR (voice recognition) (“keyword detection by the device prior to receipt of the input at the transceiver”); however, if local VR of handset 100 fails to decode the input string, the features are transmitted to base station 110 for recognition by remote VR (column 8, lines 46 to 56; column 9, lines 6 to 20: Figure 5); implicitly, the words to be decoded for these small number of simple, special voiced commands are “keywords”;

“a memory configured to store an acoustic model of the input; and a processing module coupled to the transceiver and configured to: perform speech recognition on the received input, the speech recognition comprising: recognition based on a previously

stored acoustic model, and recognition based on keyword detection in order to recognize a command" – remote VR recognizes regular voiced commands with a larger vocabulary table at remote word decoder of base station 110 (column 8, lines 28 to 56: Figure 5); implicitly, words recognized by remote word decoder for a regular voiced command involves "keyword detection"; acoustic pattern matching in a word decoder requires a mathematical model to describe the speaker's phonological and acoustic-phonetic variations for acoustic pattern matching (column 2, lines 31 to 40); acoustic pattern matching in a word decoder can be based on hidden Markov models (HMM's) (column 4, lines 13 to 21); speech signals are provided to acoustic processor 52, which requires an acoustic feature sequence as input for both recognition and training tasks (column 6, lines 62 to 67: Figure 3); thus, acoustic pattern matching by remote word decoder 114 ("a processing module") involves a stored "acoustic model" to perform speech recognition by matching an acoustic feature sequence to a stored mathematical model of a speaker's phonological and acoustic-phonetic variations;

"wherein the transceiver is further configured to transmit data to the device responsive to the command, to enable the device to provide the data in an output response" – at central communications center 42, an action signal is provided to transmitter 50, so that estimated words or a command signal are transmitted to portable phone 40 ("to transmit data to the device responsive to the command"); at portable phone 40, the estimated words or command signals are received, and then provided to control element 38; in response to the received command signal or estimated words, control element 38 provides the intended response; an intended response can be

providing information on a display screen ("to enable the device to provide the data in an output response") (column 5, lines 44 to 65: Figure 2).

Concerning independent claims 59, 81, and 83, the only elements omitted by *Jacobs et al.* are that the processing module is configured to "update or modify the keyword detection based on words within the additional input, and update the previously stored acoustic model based on the additional input". *Jacobs et al.* reasonably discloses "an acoustic model" for acoustic pattern matching by a word decoder, and keywords for special voiced commands, and even suggests training of the acoustic model. (See Column 6, Line 64 to 67) However, *Jacobs et al.* does not disclose updating or modifying keyword detection or updating a previously stored acoustic model with additional input. Still, it is known to update both acoustic models and keyword grammars during speech recognition so as to improve recognition performance by adaptation.

Concerning claims 59, 81, and 83, specifically, *Miyazawa et al.* teaches a speech recognition method for a speech interactive device, where an initial word enrollment is followed by additional word enrollment that creates standard patterns that are speaker-adapted and stored for speaker specific word enrollment. (Abstract; Column 3, Lines 39 to 48) Pre-registered words can be speaker-adapted to permit more accurate and quicker recognition, and to allow specific speakers to enroll new words suited to the user's individual needs and tastes which are not included in the non-speaker specific word registry storage. (Column 4, Lines 43 to 60) *Miyazawa et al.* refers to these words as "keywords" for keyword-spotting processing technology, and keywords include

"time", "tomorrow", and "weather" for responding to commands for information about the weather. (Column 7, Lines 50 to 57; Column 8, Lines 7 to 11; Column 9, Lines 21 to 51) Keywords are stored in the form of patterns in standard pattern memory unit 31 for the predetermined word registry. (Column 8, Lines 7 to 14) Word enrollment 81 creates standard patterns for the input voice as standard characteristic voice data, and the standard pattern is stored in standard pattern memory unit 82. (Column 10, Lines 26 to 30) Here, a standard pattern is equivalent to "an acoustic model". Thus, *Miyazawa et al.* teaches both updating or modifying keyword detection by additional word enrollment, and updating a stored acoustic model when a standard pattern is stored for speaker-specific word registration. Objectives include accommodating a wider range of conversation responses and detected phrases on an as needed basis. (Column 3, Lines 2 to 5) It would have been obvious to one having ordinary skill in the art to update and modify keyword detection and update acoustic models as taught by *Miyazawa et al.* in a distributed voice recognition system and method of *Jacobs et al.* for a purpose of accommodating a wider range of conversation responses and detected phrases as needed.

Concerning independent claims 66 and 73, *Jacobs et al.* discloses a distributed voice recognition system and method, comprising:

"receiving an audio input from a device over a network, the audio input based on speech input, wherein the audio input is the result of preliminary signal processing comprising keyword detection by the device prior to receipt of the audio input" – central

communications center or base station 42 receives speech features ("audio input") transmitted over a wireless communication network ("a network") from portable phone 40 ("the device"); a speech signal received at microphone 20 of portable phone 40 is provided to feature extraction element 22, which extracts relevant characteristics of the input speech ("preliminary signal processing") (column 5, lines 21 to 56: Figure 2); in one embodiment, handset 100 recognizes a small number of simple, special voiced commands by local VR (voice recognition) ("keyword detection by the device prior to receipt of the input at the transceiver"); however, if local VR of handset 100 fails to decode the input string, the features are transmitted to base station 110 for recognition by remote VR (column 8, lines 46 to 56; column 9, lines 6 to 20: Figure 5); implicitly, the words to be decoded for these small number of simple, special voiced commands are "keywords";

"storing an acoustic model of the audio input; performing speech recognition on the received audio input, the speech recognition comprising recognition based on a previously stored acoustic model and recognition based on keyword detection in order to recognize a command" – remote VR recognizes regular voiced commands with a larger vocabulary table at remote word decoder of base station 110 (column 8, lines 28 to 56: Figure 5); implicitly, words recognized by remote word decoder for a regular voiced command involves "keyword detection"; acoustic pattern matching in a word decoder requires a mathematical model to describe the speaker's phonological and acoustic-phonetic variations for acoustic pattern matching (column 2, lines 31 to 40); acoustic pattern matching in a word decoder can be based on hidden Markov models

(HMM's) (column 4, lines 13 to 21); speech signals are provided to acoustic processor 52, which requires an acoustic feature sequence as input for both recognition and training tasks (column 6, lines 62 to 67: Figure 3); thus, acoustic pattern matching by remote word decoder 114 involves a stored "acoustic model" to perform speech recognition by matching an acoustic feature sequence to a stored mathematical model of a speaker's phonological and acoustic-phonetic variations;

"transmitting data to the device over the network, responsive to the command, to enable the device to provide the data in an output response" – at central communications center 42, an action signal is provided to transmitter 50, so that estimated words or a command signal are transmitted to portable phone 40 ("to transmit data to the device responsive to the command"); at portable phone 40, the estimated words or command signals are received, and then provided to control element 38; in response to the received command signal or estimated words, control element 38 provides the intended response; an intended response can be providing information on a display screen ("to enable the device to provide the data in an output response") (column 5, lines 44 to 65: Figure 2).

Concerning independent claims 66 and 73, the only elements omitted by *Jacobs et al.* are "updating or modifying the keyword detection based on words within the additional input, and updating the previously stored acoustic model based on the additional input". *Jacobs et al.* reasonably discloses "an acoustic model" for acoustic pattern matching by a word decoder, and keywords for special voiced commands, and even suggests training of the acoustic model. (See Column 6, Line 64 to 67) However,

Jacobs et al. does not disclose updating or modifying keyword detection or updating a previously stored acoustic model with additional input. Still, it is known to update both acoustic models and keyword grammars during speech recognition so as to improve recognition performance by adaptation.

Concerning claims 66 and 73, specifically, *Miyazawa et al.* teaches a speech recognition method for a speech interactive device, where an initial word enrollment is followed by additional word enrollment that creates standard patterns that are speaker-adapted and stored for speaker specific word enrollment. (Abstract; Column 3, Lines 39 to 48) Pre-registered words can be speaker-adapted to permit more accurate and quicker recognition, and to allow specific speakers to enroll new words suited to the user's individual needs and tastes which are not included in the non-speaker specific word registry storage. (Column 4, Lines 43 to 60) *Miyazawa et al.* refers to these words as "keywords" for keyword-spotting processing technology, and keywords include "time", "tomorrow", and "weather" for responding to commands for information about the weather. (Column 7, Lines 50 to 57; Column 8, Lines 7 to 11; Column 9, Lines 21 to 51) Keywords are stored in the form of patterns in standard pattern memory unit 31 for the predetermined word registry. (Column 8, Lines 7 to 14) Word enrollment 81 creates standard patterns for the input voice as standard characteristic voice data, and the standard pattern is stored in standard pattern memory unit 82. (Column 10, Lines 26 to 30) Here, a standard pattern is equivalent to "an acoustic model". Thus, *Miyazawa et al.* teaches both updating or modifying keyword detection by additional word enrollment, and updating a stored acoustic model when a standard pattern is stored for speaker-

specific word registration. Objectives include accommodating a wider range of conversation responses and detected phrases on an as needed basis. (Column 3, Lines 2 to 5) It would have been obvious to one having ordinary skill in the art to update and modify keyword detection and update acoustic models as taught by *Miyazawa et al.* in a distributed voice recognition system and method of *Jacobs et al.* for a purpose of accommodating a wider range of conversation responses and detected phrases as needed.

Concerning claims 63, 70, and 77, *Jacobs et al.* discloses that portable phone 40 may receive a command signal or estimated words, and control element 38 provides an intended response; the intended response may be to provide information to display screen on the portable phone (column 5, lines 62 to 65: Figure 2); thus, the response will be "a text message" of information on a display of portable phone 40.

Concerning claims 64, 71, and 78, *Jacobs et al.* discloses that features are provided to local word decoder 106 which searches its small vocabulary to recognize the input speech; if local word decoder 106 fails to decode the input string and determines that remote VR should decode it, the features are transmitted to remote word decoder 110 (column 9, lines 6 to 15: Figure 5); thus, handset 100 will transmit the speech features ("the input") for remote VR when the input string at handset ("the device") "is not capable of being processed by the device".

7. Claims 61 to 62, 65, 68 to 69, 72, 75 to 76, and 79 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Jacobs et al.* in view of *Miyazawa et al.* as applied to claims 59, 66, and 73 above, and further in view of *Houser et al.*

Concerning claims 61 to 62, 68 to 69, and 75 to 76, *Jacobs et al.* discloses sending information for display on a display screen of portable phone. (Column 5, Lines 62 to 65: Figure 2) Implicitly, the displayed information is text. However, *Jacobs et al.* does not disclose providing audio data and video data as data that is a response to a user command, although audio data might be suggested in an environment directed to a wireless portable phone. However, it is known for speech commands to access information in a variety of forms on cellular phones, where the information includes music and video. Specifically, *Houser et al.* teaches a speech interface for controlling a device such as a television and for controlling access to broadcast information such as video, audio, and/or text information in accordance with recognized utterances of a user. (Abstract) An objective is to afford ease of use as well as permitting the implementation of tasks which are not easily implemented using menu screens and key presses. (Column 2, Lines 23 to 29) It would have been obvious to one having ordinary skill in the art to provide audio data and video data to a user in response to a user's voice command as taught by *Houser et al.* in a distributed voice recognition system and method of *Jacobs et al.* for a purpose of permitting implementation of tasks which would be difficult to perform using menu screens and key presses.

Concerning claims 65, 72, and 79, *Houser et al.* discloses that information is retrieved from an information distribution center 12 in response to commands from terminal unit 16 for accessing information (column 5, line 39 to column 6, line 14: Figure 1); additionally, electronic programming guide (EPG) data is accessed from an information provider 114-3, including television schedule information arranged by time and channel, and transmitted to subscriber units (column 22, line 19 to 51: Figure 2C).

Response to Arguments

8. Applicants' arguments filed 12 January 2010 have been considered but are moot in view of the new grounds of rejection, necessitated by amendment.

Applicants have substantially rewritten independent claims 59, 66, 73, 81, and 83, requiring a new search and new grounds of rejection set forth herein. Accordingly, Applicants' arguments are moot.

Conclusion

9. Applicants' amendment necessitated the new grounds of rejection presented in this Office Action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicants are reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (571) 272-7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic

Art Unit: 2626

Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Martin Lerner/
Primary Examiner
Art Unit 2626
February 5, 2010